

**CONSTRUCT VALIDITY OF CRITICAL THINKING SKILL INSTRUMENT
IN TEACHING MATHEMATIC IN JUNIOR HIGH SCHOOL****R. Rosnawati, Badrun Kartowagiran***Yogyakarta State University
Indonesia***Abstract**

This study aims to evaluate the construct validity of cognitive instruments of critical thinking in teaching mathematics in junior high school. The instrument was developed using a method that allows to design an obvious question to measure both the content and critical thinking skills. Critical thinking instrument consists of 42 multiple choice questions with four alternative choices. The study involved 656 students from junior high school in Yogyakarta. The findings indicate that critical thinking has satisfactory construct validity by five factors extracted and confirmed by confirmatory factor analysis. CFA revealed the model fit. (CFI = 0.97; RMSEA = 0.059; IFI = .0.97; PNFI = 0.90, and $\chi^2/df = 3.10$)

Keywords: construct validity, critical thinking

INTRODUCTION

In fact, there are different definitions stronghold against critical thinking, conceptual diversity comes from the fact that critical thinking is studied in different fields and applied in various contexts. The existence of different views on critical thinking, affect the approach to teaching critical thinking skills, which are generally divided into two different views are referred to as generalist view and non-generalist. The generalist maintain that no generalized thinking skills that can be taught without involving a particular context. Non-generalist position championed by John McPeck (1981) maintain that thinking can only occur with reference to some specific contexts. It affects the way to measure students' thinking skills. (Mason, 2007)

There is no consensus about how critical thinking should be measured. Results of previous studies show there are a number of instruments of critical thinking among others: Watson-Glaser Critical Thinking Appraisal Skills (WGCTA, Watson & Glaser, 1980), Ennis-Weir Critical Thinking Essay Test (EWCTET, Ennis & Weir, 1985), Cornell Critical Thinking Test (CCTT, Ennis, Millman, & Tomko, 1985), the California Critical Thinking Skills Test (CCTST, Facione, 1990b), and Halpen Critical Thinking Using Everyday Situation Assessment (HCTAES, Halpern, 2007). (Ku, 2009: 71)

WGCTA, CCTST, CCTST, and CCTDI are examples of instruments that utilize a single multiple-choice response format. WGCTA, devided of three forms, the standard form of WGCTA consists of 80 items that measure the ability of the five aspects of critical thinking: inference, recognition of assumptions, deductions, interpretation, and evaluation of arguments. CCTST consists of two forms, each of which consists of 34 items that measure five aspects of critical thinking: interpretation, analysis, evaluation, inference, and explanation. CCTT consists of consists of two levels (ie, X and Z), in the form of multiple-choice-based story. Level X contains 71 items designed for grade 4 to students, and the rate of Z contains 62 items that will be used for high school and college students. Overall, the two forms of CCTT measure critical thinking are seven aspects of induction, deduction, credibility, assumptions, semantics, definition, and prediction. EWCTET is a test of critical thinking skills in the form of essays

open (open-ended) are intended for high school or college students. (Adams, Whitlow, Stover, & Johnson, 1996).

HCTAES is an example of critical thinking instrument which has the form of multiple-choice and open-ended. HCTAES measure critical thinking skills using the questions in the context of an authentic and believable. The test consists of 25 open-ended questions and 50 multiple choice questions. There is evidence that the responses of multiple choice and open-ended essay measure cognitive abilities separately (Bridgeman & Moran, 1996; Halpern, 2007). Form of open-ended on the part of HCTAES try to reveal more of the components of dispositional thinking, as it allows candidates to indicate whether they are likely to apply appropriate skills, while multiple-choice format to measure the "recognition memory" (Halpern, 2007). Questions from HCTAES representing five skill categories are: verbal reasoning (eg, recognizes the widespread use of language or misleading), argument analysis (for example, recognizes the reasons and conclusion of the argument), hypothesis testing, using likelihood and uncertainty choose between alternative solutions), as well as decision making and problem solving (eg, identify goals problems, generate and choose between alternative solutions)

The instrument was developed in the previous exposure is an instrument that measures critical thinking skills in general. Very little is associated with a particular subject. The purpose of this research is to develop critical thinking instruments associated with the topic algebra. The specific objective of this research is to identify attributes/aspects that can be used to measure critical thinking skills. This study is necessary because one of the objectives to be achieved in teaching junior high school mathematics is achieving critical thinking skills (Permendikbud No 68, 2013).

To be able to formulate mathematical learning objectives, content, and assessment of critical thinking in mathematics is required operational definition of critical thinking in mathematics. Bloom's Taxonomy developed four classification of learning objectives, namely cognitive, affective, psikomotor, and perceptual. Generally, developed only three aspects: cognitive, affective and psikomotor. Cognitive domains related to intellectual results concerning two aspects: cognitive and intellectual abilities. Learning objectives compiled from the simple to the complex, ie starting from remembering, understanding, application, analysis, synthesis, and evaluation. Learning objectives at the level of recall and understand not require critical thinking skills, but the level of application, analysis, synthesis and evaluation requires thinking skills that are the hallmark of critical thinking. However, four of these levels can not measure critical thinking disposition.

According mathematics's curriculum to junior high school in Indonesia, algebra consist of number, the algebra operations, linear equations and inequalities, relations and functions and systems of linear equations. This is in accordance with standards established by the NCTM (2000: 37) that learning algebra of class early to grade 12 should enable all students to understand patterns, relations and functions, representation and analyze situations and structures using algebraic symbols of mathematics, using mathematical models to represent and understand quantitative relationships, and analyze changes in a variety of contexts. Furthermore, Lawrence and Hennecy (2002: xi) states think algebra includes a set of understanding necessary to interpret the world by translating information or events into mathematical language to explain and predict phenomena, leading to abstract thinking needed to succeed in understanding algebra, through analyzing topics changes, functions, variables, set.

Furthermore Kaput (1998: 26) defines an algebraic reasoning as a complex composition of the five forms that interact conceptually. The students experience with this type of reasoning should start early associated with:

1. Algebra as a generalization and formalizing patterns and constraints, but not exclusively, as a generalization of algebraic reasoning arithmetic and algebra as a generalization of quantitative reasoning.
2. Algebra as syntactic manipulation - guided formalism
3. Algebra as the study of the structure and system of calculation and relations abstracted.

4. Algebra as the study of functions, relationships, and the combined variations
5. Algebra as a cluster of (a) modeling and (b) the phenomenon of language-controller

From the definition above in this study will be used topics related to arithmetic and algebra that consists of pattern, shape algebra operations, sets, relations and functions, as well as the system of linear equations.

Based on the study of critical thinking and the content of the above algebra, in this study identified conceptualization or type of critical thinking in algebra topics, namely:

1. Using inductive algebra problems
2. Using deductive on algebra problems
3. Provide examples of related concepts in algebra
4. Checking arguments or statements related to algebra
5. Provide an assessment of two statements or concepts in algebra
6. Using the concept to draw conclusions and make generalizations
7. Analyze the application of a concept in algebra

Seventh kind of thinking is poured into the lattice problem, and further develop appropriate assessment rubric with critical thinking skills are developed..

RESEARCH METHOD

The participants of this study involving 656 students from junior high school junior high school in Yogyakarta. In this study, an instrument has been validated by a expert who are lecturers in the field of assessment, mathematic educations, psychology. The quality of the instrument is also seen from the empirical validity and reliability. Reliability is a measure of the internal consistency of indicators a construct that indicates the degree of the extent to which each indicator indicates a latent constructs.

Two statistical procedures used in analyzing the data. First, Bartlett Test of Sphericity by using SPSS. The statistical test used to determine whether there is any correlation between variables. Second confirmatory factor analysis (CFA) using LISREL program. CFA is used to test the measurement model. The first phase included all indicators at each construct or variable, then performed the modification of indicators and indices to obtain a fit model (final stage). The next phase of the endogenous constructs confirmatory factor analysis using confirmatory factor analysis measurement model.

Confirmatory Factor Analysis (CFA) is conducted to estimate factor loading of variables. In fact, a factor loading presents the level of a regression path from a latent to its indicators. In this study, all of latent variables had at least three indicators (the questionnaire item). According to (Hair et al., 2010), an acceptable factor loading value is more than 0.5 and when it is equal to 0.7 and above it is considered good for one indicator. The level of CR is another guideline to review convergent validity. According to (Hair et al., 2010), the acceptable value of CR is 0.7 and above. It is calculated by Equation 1.

$$CR = \frac{(\sum_{i=1}^n \lambda_{xi})^2}{(\sum_{i=1}^n \lambda_{xi})^2 + (\sum_{i=1}^p (\varepsilon_i))}$$

With $\varepsilon_i = 1 - (\lambda_{xi})^2$

Some of the indicators used in assessing the fit model, among others, the Non-normed Fit Index (NNFI), Tucker-Lewis Index (TLI), Comparative Fit Index (CFI), and Standard Root Mean Square, Residual (RMSEA). A series of simulation studies by Hu and Bentler (1999) states that the cutoff value close to 0.08 for the RMSEA and close to the cutoff value of 0.95 which supports a suitable model in relation to the data. In addition to these indicators, χ^2/df is used to reduce the sensitivity of the test χ^2 with sample size. As a rule of thumb, the value χ^2/df acceptable as it should be below 5 and preferably below 3 (Hu & Bentler, 1999).

To compare between several alternative models, use the value of the AIC. Size suitability parsimony (Parsimonious Fit Measures, PFM) inform simplicity the model in relation to the number of estimate parameters, in terms of parsimonious normed fit index (PNFI). Confirmatory factor, changes χ^2 calculated to assess the extent to which alternative models fitted the data better (Kline 2005). When the model is not nested, the index Akaike Information Criterion (AIC) is used. Lower AIC value support non-nested models (Kline 2005).

DISCUSSION

Development of an instrument to measure critical thinking skills in mathematics learning is done through a number of stages that can be divided into two main phases, namely: a qualitative and quantitative phases. Stages involve qualitative focus group discussion (FGD) to get aspects of critical thinking skills, while the quantitative stage involves confirmatory factor analysis (CFA) to test the suitability of the measurement model, reliability and validity of the attributes of critical thinking skills in mathematics.

Product mathematical instruments critical thinking based on the input of each expert and reliability test results are then used to obtain the empirical validation of the 656 students from 4 junior high schools in DIY. Empirical validity test serves to determine the accuracy and precision of measurement of a measuring instrument for measuring a variable in accordance with the concept to be measured. The higher the index value of the validity of a measuring instrument, the more precise is also a measuring instrument in achieving the target.

The initial step prior to testing using EFA, is testing KMO and Bartlett's Test useful to determine the feasibility of each item to be tested. Testing instrument critical thinking done on all indicators simultaneously, and the results SPSS display shows the value of KMO MSA obtained for $0.936 > 0.05$ with $p\text{-value } 0.000 < 0.05$. The test results lead to the conclusion that the data correlation matrix has enough so that it can be used to perform factor analysis. Further analysis of the value of the MSA item shows all items have a value above 0.3 MSA is within acceptable limits, which means that all items can be included in the subsequent analysis.

First-order CFA testing done by making the measurement model to describe how well the indicators in measuring devices can be used as a latent variable measuring instrument in this critical thinking. In this test the model one latent variable critical thinking by making the 65 items as indicators measured directly. This test will prove that all measurement items measure a single variable. The test results of the measurement model of first-order CFA produces $p\text{-value} = 0.00 (< 0.05)$ and $RMSEA = 0.096$, which means that the model does not fit with the data. Further modification of the model, by eliminating some of the items that contribute little to the measuring instrument. There are 42 items that were passed test results first order CFA, of 42 items that have passed the test the first-order CFA then back through the process of testing first-order CFA. The test results of the first-order CFA measurement model with 42 items by adding error covariance produce $p\text{-value} = 0.09068 (p > 0.05)$ and $RMSEA = 0.049 < 0.08$, so that concluded this model really fit with the data. Validity is indicated by the value of the loading factor and t values greater than 1.96 for all items

Next step is testing the measurement model of critical thinking in teaching mathematics. Determining the number of factors/aspects in measuring students' critical thinking skills in teaching mathematics. The initial step determining the number of factors is based on the eigenvalues of matrix data transformation results. In this study eigenvalues greater than 1 as 13. So, the number of factors that can be set up to 13. Furthermore, this study will look for a model that is dependent on a number of independent factors that best fits the data. In matrix

components, interpretation of factors is difficult due to the many variables that have some similar loading factor is at several different factors, so that a rotation of the matrix components with varimax rotation method. From the results obtained varimax rotation using the data classification based on various factors independently. Four kinds of models are presented in Table 1.

Table 1. Number of Items

Model	Number of Factor	Number of items for each model								Σ
		1	2	3	4	5	6	7	8	
Model 1	8 Factor	8	6	8	4	7	4	2	3	42
Model 2	7 Factor	7	8	5	8	8	4	2		42
Model 3	6 Factor	5	6	9	7	8	7			42
Model 4	5 Factor	11	7	7	8	9				42

Here are the test results of the CFA for some models as listed in Table 1 are briefly presented in Table 2 below.

Table 2 Results of the CFA Test Measurement Model of Critical Thinking Instrument

No	Model	χ^2	df	χ^2/df	RMSE	AIC	PGFI	PNFI
1	8 factor independent	2746.1	791	3.47	0.061	2970.00	0.73	0.87
2	7 factor independent	2584.08	719	3.59	0.063	2786.43	0.73	0.88
3	6 factor independent	3038.98	804	3.78	0.065	3236.98	0.73	0.88
4	5 factor independent	2675.53	809	3.31	0.059	2806.00	0.74	0.90

From the test results of the CFA for each model as shown in the Table 2, the model 4 with a lot of factors is independent as much as 5 has a value size suitability parsimony normed Chi-Square or CMIN/DF of 3.31, the size of the suitability of parsimony (Parsimonius Fit Measures, PFM) inform simplicity model in relation to the number of parameters diestimas, in terms of parsimonious normed fit index (PNFI) larger than the other models.

Furthermore, from the value of the construct reliability (CR) for each factor in each model is presented in Table 3. In model 4 all factors have a value of $CR \geq 0.70$, while the other models consist the factor with a value of $CR \geq 0.70$. Based on that in this study the model 4 is the best measurement model that best fits the data.

Based upon the results of the measurement model, a model that fits the data is the model 4 is formed 5 factors/aspects of critical thinking in learning mathematics. Path diagram can be seen in the appendix. Factors that are formed are as follows: reasoning, interpretation, analysis, evaluation, and inference.

Table 3 Value Construct Reliability (CR)

Model	CR	Factor							
		1	2	3	4	5	6	7	8
1	Σ item	8	6	8	4	7	4	2	3
	$CR \geq 0.70$	0.89	0.79	0.77	0.79	0.82	0.47	0.50	0.51
2	Σ item	7	8	5	8	6	4	2	
	$CR \geq 0.70$	0.86	0.82	0.79	0.85	0.83	0.62	0.50	
3	Σ item	5	6	9	7	8	7		
	$CR \geq 0.70$	0.84	0.82	0.82	0.80	0.86	0.68		
4	Σ item	11	7	7	8	9			
	$CR \geq 0.70$	0.88	0.79	0.84	0.88	0.86			

CONCLUSIONS AND SUGGESTIONS

Factors of critical thinking in mathematics is reasoning, interpretation, analysis, evaluation, and inference. From the results of the model test CR value for each aspect are 0.88, 0.79, 0.84, 0.88, and 0.86 which are in good category.

It's important to develop a good instrument and in accordance with the learning objectives to be achieved. One of the learning objectives to be achieved is critical thinking skills. Furthermore, when it has obtained information from students about the weakness of critical thinking in more detail, if specific information obtained may indicate a weak aspect of critical thinking, then the information can be used to make improvements under the conditions of the students so that the students' critical thinking skills can be improved.

REFERENCES

- Hair, JF, Black, W, Babin, B.J, Anderson R E. (2010). *Multivariate Data Analysis*. (7th Ed.), NJ: Pearson-Prentice Hall, Upper Saddle River,
- Hu, Li-tze and Bentler, PM. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives, *Structural Equation Modeling: A Multidisciplinary Journal*, 6:1, 1-55
- Kaput, J. J. (1995). A research base supporting long term algebra reform. In D. T. Owens, M. K. Reed, & G. M. Millsaps (Eds.), *Proceedings of the seventeenth annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 71-94). Columbus, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.
- Kline, R.B. (2011). *Principles and practice of structural equation modeling*. New York: The Guilford Press
- Ku, K.Y.L. (2009). Assessing students' critical thinking performance: Urging for measurements using multi-response format. *Thinking Skills and Creativity*. 4 , 70–76.
- Lawrence, A & Hennessy, C. (2002). *Lessons for algebraic thinking*. Sausalito, CA : Math Solutions Publications
- Mason, M. (2007). Critical thinking and learning. *Educational Philosophy and Theory*. DOI:

10.1111/j.1469-5812.2007.00343.x

National Council of Teachers of Mathematics. (2009). *Focus in high school mathematics reasoning and sense making*. Reston: National Council of Teachers of Mathematics.

O'Hare, L. O., & McGuinness, C. (2009). Measuring critical thinking, intelligence, and academic performance in psychology undergraduates. *The Irish Journal of Psychology*, 30(3–4), 123–131.

Thompson, B. & Vacha-Haase, T. (2000) Psychometrics is datametrics: the test is not reliable. *Educational and Psychological Measurement*. 60; 174. DOI: 10.1177/0013164400602002

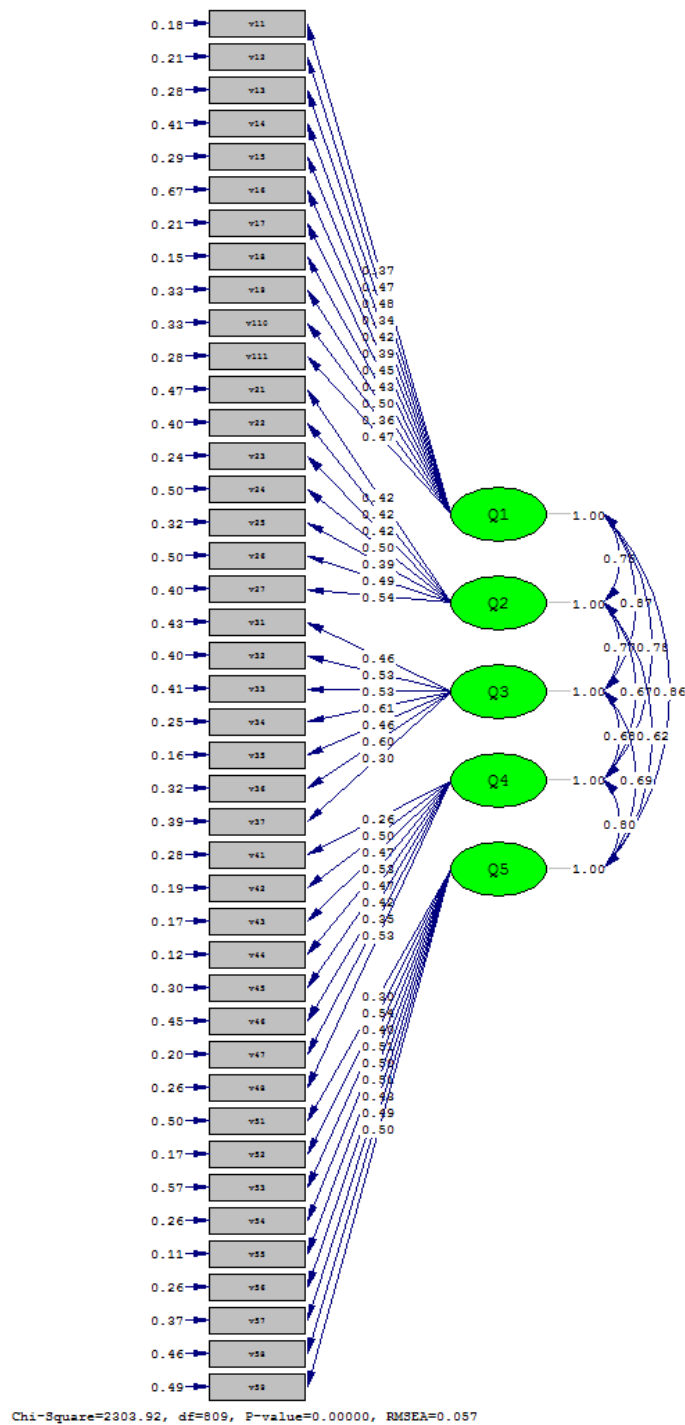


Figure 1. measurement model